

A practical attack to Bouftass's cryptosystem

Yang Zhang

May 4, 2016

Texas Tech University
Lubbock, TX, United States
yang22.zhang@ttu.edu

Abstract

Recently, a new fast public key exchange protocol was presented by S. Bouftass. The protocol is based on the difficulty of inverting the function $F(x) = \lfloor (zx \bmod 2^p)/2^q \rfloor$. In this paper, we describe a practical attack against this protocol based on Closest Vector Problem (CVP) and Gaussian lattice reduction.

Keywords public key exchange cryptoanalysis CVP Gaussian lattice reduction

1 Introduction

In public key cryptography, the security of traditional methods is based on number theoretic problems, and suffers from high computational cost due to problems such as dealing with large numbers. Each user in a public key system has a pair of cryptographic keys, consisting a public key and a private key. These are related through a hard mathematical inversion problem, so that the private key cannot be feasibly derived from the public key. A standard implementation of public key cryptography is based on the Diffie-Hellman key agreement protocol[1]. This protocol allows two users, Alice and Bob, to exchange a secret key over an insecure communication channel. It can be described as following:

1. Alice and Bob openly agree upon a large prime p and $g \in \mathbb{Z}_p^*$.
2. Alice randomly chooses the secret integer $a \in [1, p-1]$.
3. Alice computes $A = g^a \bmod p$, and publishes A .
4. Bob randomly chooses the secret integer $b \in [1, p-1]$,
5. Bob computes $B = g^b \bmod p$, and publishes B .
6. Alice computes the secret integer $K_A = B^a \bmod p = g^{ba} \bmod p$.
7. Bob computes the secret integer $K_B = A^b \bmod p = g^{ab} \bmod p$.

Then Alice and Bob can get the same shared secret key $K = K_A = K_B$. The eavesdropper Eve knows p, g, A and B , and she needs to compute the secret key K . For this, it suffices to solve one of the discrete logarithm problems:

$$A = g^a \pmod{p} \quad \text{and} \quad B = g^b \pmod{p}$$

for the unknowns a or b . If p is a very large prime of say 2048 bits, then the problem becomes computationally hard, and it is considered infeasible. For maximum security p should be a safe prime, i.e. $(p-1)/2$ is also a prime, and g a primitive root of p [2].

Recently, to construct a cryptosystem which is not based on number theory, S.Bouftass described a new public key exchange protocol relying on the difficulty of inverting the function $F(x) = \lfloor (zx \bmod 2^p)/2^q \rfloor$ [3]. In our work, we find that this system is not secure, we can easily break this system based on the closest vector problem ([4],[5]) and Gaussian lattice reduction[4]. This paper is organized as follows, in section 2 we give a general description of S.Bouftass's new protocol; section 3 gives our method to break this system and an example; The last section is conclusion.

2 S.Bouftass's new public key exchange cryptosystem

Throughout, if n is an integer and $s \in \mathbb{N}$, we use $a \bmod n$ to denote the nonnegative remainder of a divided by n . We will use the same notation as in [3] to exchange the secret key. Alice and Bob should agree on some integers: l, m, p, q, r, z , where z is l bits long, $p+q = l+m$, $p > m+q+r$, and $r > 128$. The protocol is then described as follows,

1. Alice and Bob agree upon the integers l, m, p, q, r, z . Alice randomly selects a private m bit positive integer x , and Bob selects a private m bit positive integer y .
2. Alice computes $U = \left\lfloor \frac{(xz) \bmod 2^p}{2^q} \right\rfloor$ and sends it to Bob.
3. Bob computes $V = \left\lfloor \frac{(yz) \bmod 2^p}{2^q} \right\rfloor$ and sends it to Alice.
4. Alice computes $W_a = \left\lfloor \frac{(xV) \bmod 2^{p-q}}{2^{r+m}} \right\rfloor$.
5. Bob computes $W_b = \left\lfloor \frac{(yU) \bmod 2^{p-q}}{2^{r+m}} \right\rfloor$.
6. The shared secret key is $K = W_a = W_b$ when $r > 128$.

3 Practical attack to this cryptosystem

Let l, m, p, q, r and z be fixed as above and let $F(x) = \left\lfloor \frac{(xz) \bmod 2^p}{2^q} \right\rfloor = u$, then we have

$$2^q u + y = xz \pmod{2^p}$$

for some integer y with $0 \leq y < 2^q$, i.e.

$$xz \equiv 2^q u + y \pmod{2^p}. \tag{1}$$

Hence, finding an element $x \in F^{-1}(\{u\})$ is equivalent to finding a proper vector $\begin{bmatrix} x \\ y \end{bmatrix}$ that satisfies equation (1), and $0 \leq y < 2^q$.

Theorem 1. All solutions to equation (1) are of the form $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_0 + \alpha_1 x_1 + \alpha_2 x_2 \\ y_0 + \alpha_1 y_1 + \alpha_2 y_2 \end{bmatrix}$, where $\alpha_1, \alpha_2 \in \mathbb{Z}$ and

$$\begin{aligned} x_0 &= \left\lfloor \frac{2^q u}{z} \right\rfloor, & y_0 &= zx_0 - 2^q u; \\ x_1 &= \left\lfloor \frac{2^q u}{z} \right\rfloor, & y_1 &= zx_1 - 2^p; \\ x_2 &= \left\lfloor \frac{2^q u}{z} \right\rfloor + 1, & y_2 &= zx_2 - 2^p. \end{aligned}$$

Proof. Let x_0, y_0, x_1, y_1 and x_2, y_2 be the values that are defined above, then it is obvious that for all integers α_1 and α_2 we always have

$$(x_0 + \alpha_1 x_1 + \alpha_2 x_2)z \equiv 2^q u + (y_0 + \alpha_1 y_1 + \alpha_2 y_2) \pmod{2^p},$$

i.e. all vectors of the form $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_0 + \alpha_1 x_1 + \alpha_2 x_2 \\ y_0 + \alpha_1 y_1 + \alpha_2 y_2 \end{bmatrix}$ are solutions to equation (1) for $\forall \alpha_1, \alpha_2 \in \mathbb{Z}$.

On the other hand, since

$$x_0 z \equiv 2^q u + y_0 \pmod{2^p},$$

let $\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix}$ be an arbitrary solution to equation (1), then

$$(\hat{x} - x_0)z \equiv \hat{y} - y_0 \pmod{2^p}.$$

Hence there exists $n \in \mathbb{Z}$, such that

$$(\hat{x} - x_0)z = n2^p + \hat{y} - y_0, \tag{2}$$

i.e.

$$\hat{x} - x_0 = \frac{n2^p + \hat{y} - y_0}{z} = n \left\lfloor \frac{2^p}{z} \right\rfloor + C + \frac{\hat{y} - y_0}{z}$$

for some number $C \in \mathbb{R}$.

Since $\hat{x} - x_0$ is integer, $C + \frac{\hat{y} - y_0}{z}$ should also be an integer, call it N . Now we have

$$\hat{x} - x_0 = n \left\lfloor \frac{2^p}{z} \right\rfloor + N = (n - N) \left\lfloor \frac{2^p}{z} \right\rfloor + N \left(\left\lfloor \frac{2^p}{z} \right\rfloor + 1 \right).$$

Let $n - N = \alpha_1$ and $N = \alpha_2$, we can get

$$\hat{x} - x_0 = \alpha_1 x_1 + \alpha_2 x_2. \tag{3}$$

Next, combining equations (2) and (3) we have

$$\begin{aligned} \hat{y} - y_0 &= (\alpha_1 x_1 + \alpha_2 x_2)z - n2^p \\ &= \alpha_1(zx_1 - 2^p) + \alpha_2(zx_2 - 2^p) \\ &= \alpha_1 y_1 + \alpha_2 y_2. \end{aligned}$$

So $\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} x_0 + \alpha_1 x_1 + \alpha_2 x_2 \\ y_0 + \alpha_1 y_1 + \alpha_2 y_2 \end{bmatrix}.$

□

1. Set $\mathbf{u}_1 \leftarrow \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$, $\mathbf{u}_2 \leftarrow \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$ and $\text{done} \leftarrow 0$;

2. While $\text{done} = 0$, do

- $c_1 \leftarrow \text{Round}\left(\frac{\langle \mathbf{u}_1, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}\right)$; $\mathbf{u}_1 \leftarrow \mathbf{u}_1 - c_1 \mathbf{u}_2$;
- $c_2 \leftarrow \text{Round}\left(\frac{\langle \mathbf{u}_1, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\right)$; $\mathbf{u}_2 \leftarrow \mathbf{u}_2 - c_2 \mathbf{u}_1$;
- if $c_1 = 0$ and $c_2 = 0$, then $\text{done} \leftarrow 1$.

3. Solve the equation $[\mathbf{u}_1, \mathbf{u}_2] \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$

4. $a_1 \leftarrow \lfloor \alpha_1 \rfloor$, $a_2 \leftarrow \lfloor \alpha_2 \rfloor$;

5. $\begin{bmatrix} x \\ y \end{bmatrix} \leftarrow \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} - a_1 \mathbf{u}_1 - a_2 \mathbf{u}_2$

Note: In this algorithm, we let $\text{Round}(\pm \frac{1}{2}) = 0$.

Theorem 2. *The following algorithm can be used to find a minimal solution $\begin{bmatrix} x \\ y \end{bmatrix}$ of equation (1), with respect to the norm induced by an arbitrary inner product $\langle -, - \rangle$ on \mathbb{R}^2 .*

Proof. First we show that the algorithm terminates.

Without loss of generality, we can assume $c_1 = \text{Round}\left(\frac{\langle \mathbf{u}_1, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle}\right) \neq 0$, and let $\frac{\langle \mathbf{u}_1, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle} = c_1 + \varepsilon$, where $-\frac{1}{2} \leq \varepsilon \leq \frac{1}{2}$. Then

$$\begin{aligned} \|\mathbf{u}_1 - c_1 \mathbf{u}_2\|^2 &= \|\mathbf{u}_1\|^2 + c_1^2 \|\mathbf{u}_2\|^2 - 2c_1 \langle \mathbf{u}_1, \mathbf{u}_2 \rangle \\ &= \|\mathbf{u}_1\|^2 - \|\mathbf{u}_2\|^2 \left(2c_1 \frac{\langle \mathbf{u}_1, \mathbf{u}_2 \rangle}{\langle \mathbf{u}_2, \mathbf{u}_2 \rangle} - c_1^2 \right) \\ &= \|\mathbf{u}_1\|^2 - \|\mathbf{u}_2\|^2 (2c_1 \varepsilon + c_1^2) \end{aligned}$$

Case I: Suppose $c_1 > 0$, then $c_1 \geq 1$. Since $\text{Round}(\frac{1}{2}) = 0$, we have either $-\frac{1}{2} < \varepsilon < 0$ or $\varepsilon \geq 0$. Hence $c_1 + 2\varepsilon > 0$, and $2c_1\varepsilon + c_1^2 > 0$ as well.

Case II: Suppose $c_1 < 0$, then $c_1 \leq -1$. Since $\text{Round}(-\frac{1}{2}) = 0$, we have either $0 < \varepsilon < \frac{1}{2}$ or $\varepsilon \leq 0$. Now $c_1 + 2\varepsilon < 0$, so $2c_1\varepsilon + c_1^2 > 0$.

So by both of these two cases, we always have $2c_1\varepsilon + c_1^2 > 0$, i.e. $\|\mathbf{u}_1 - c_1 \mathbf{u}_2\|^2 < \|\mathbf{u}_1\|^2$. By a similar argument we can get that $\|\mathbf{u}_2 - c_2 \mathbf{u}_1\|^2 < \|\mathbf{u}_2\|^2$. That means $\|\mathbf{u}_1\|$ and $\|\mathbf{u}_2\|$ are strictly decreasing. Since there are only finite number of elements in \mathcal{L} with norm less than $\max\left\{\left\|\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}\right\|, \left\|\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right\|\right\}$, the algorithm must terminate.

Furthermore when $c_1 = c_2 = 0$, it's trivial to see that

$$|\langle \mathbf{u}_1, \mathbf{u}_2 \rangle| \leq \frac{1}{2} \min\{|\langle \mathbf{u}_1, \mathbf{u}_1 \rangle|, |\langle \mathbf{u}_2, \mathbf{u}_2 \rangle|\}.$$

Now we show that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is a basis of \mathcal{L} . By the algorithm, it's easy to see that \mathbf{u}_1 and \mathbf{u}_2 are linear combinations of $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ and $\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$, so $\text{Span}_{\mathbb{Z}}(\mathbf{u}_1, \mathbf{u}_2) \subseteq \mathcal{L}$; on the other hand, every

step of the algorithm is invertible, we also have $\mathcal{L} \subseteq \text{Span}_{\mathbb{Z}}(\mathbf{u}_1, \mathbf{u}_2)$. Hence after terminating, $\mathcal{L} = \text{Span}_{\mathbb{Z}}(\mathbf{u}_1, \mathbf{u}_2)$.

Since $\mathbf{u}_1, \mathbf{u}_2$ are linearly independent over \mathbb{R} , there exist α_1, α_2 in \mathbb{R} , such that

$$\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2.$$

Let $a_1, a_2 \in \mathbb{N}$ with $a_1 = \text{Round}(\alpha_1)$ and $a_2 = \text{Round}(\alpha_2)$, now we want to show that $\left\| \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} - a_1 \mathbf{u}_1 - a_2 \mathbf{u}_2 \right\|$ is minimized. Let $z = b_1 \mathbf{u}_1 + b_2 \mathbf{u}_2$ be an arbitrary vector in \mathcal{L} , and let

$$d = \left\| \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} - z \right\|^2 - \left\| \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} - a_1 \mathbf{u}_1 - a_2 \mathbf{u}_2 \right\|^2, \quad (4)$$

we have the following cases:

Case 1: $a_1 = b_1, a_2 = b_2$, then $d = 0$.

Case 2: $a_1 \neq b_1, a_2 = b_2$, then

$$\begin{aligned} d &= \|(\alpha_1 - b_1) \mathbf{u}_1 + (\alpha_2 - b_2) \mathbf{u}_2\|^2 - \|(\alpha_1 - a_1) \mathbf{u}_1 + (\alpha_2 - a_2) \mathbf{u}_2\|^2 \\ &= (\alpha_1 - b_1)^2 \|\mathbf{u}_1\|^2 + (\alpha_2 - b_2)^2 \|\mathbf{u}_2\|^2 + 2(\alpha_1 - b_1)(\alpha_2 - b_2) \langle \mathbf{u}_1, \mathbf{u}_2 \rangle \\ &\quad - (\alpha_1 - a_1)^2 \|\mathbf{u}_1\|^2 - (\alpha_2 - a_2)^2 \|\mathbf{u}_2\|^2 - 2(\alpha_1 - a_1)(\alpha_2 - a_2) \langle \mathbf{u}_1, \mathbf{u}_2 \rangle \\ &= (\alpha_1 - b_1 + \alpha_1 - a_1) (a_1 - b_1) \|\mathbf{u}_1\|^2 + 2(a_1 - b_1) (\alpha_2 - a_2) \langle \mathbf{u}_1, \mathbf{u}_2 \rangle \\ &\geq 2(\alpha_1 - b_1 + \alpha_1 - a_1) (a_1 - b_1) |\langle \mathbf{u}_1, \mathbf{u}_2 \rangle| + 2(a_1 - b_1) (\alpha_2 - a_2) \langle \mathbf{u}_1, \mathbf{u}_2 \rangle. \end{aligned}$$

If $a_1 > b_1$, we have $a_1 - b_1 > 0$, $2\alpha_1 - a_1 - b_1 \geq 1 \geq |\alpha_2 - a_2|$, i.e. $d \geq 0$; if $a_1 < b_1$, then $a_1 - b_1 < 0$, $2\alpha_1 - a_1 - b_1 \leq -1$, but $|2\alpha_1 - a_1 - b_1| \geq |\alpha_2 - a_2|$, we still have $d \geq 0$.

Case 3: $a_1 = b_1, a_2 \neq b_2$, this is the same as Case 2.

Case 4: $a_1 \neq b_1, a_2 \neq b_2$, then

$$\begin{aligned} d &= \|(\alpha_1 - b_1) \mathbf{u}_1 + (\alpha_2 - b_2) \mathbf{u}_2\|^2 - \|(\alpha_1 - b_1) \mathbf{u}_1 + (\alpha_2 - a_2) \mathbf{u}_2\|^2 \\ &\quad + \|(\alpha_1 - b_1) \mathbf{u}_1 + (\alpha_2 - a_2) \mathbf{u}_2\|^2 - \|(\alpha_1 - a_1) \mathbf{u}_1 + (\alpha_2 - a_2) \mathbf{u}_2\|^2 \\ &\geq 0. \end{aligned}$$

The inequality is because of Case 2 and 3.

Above all, the norm of the vector $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} - a_1 \mathbf{u}_1 - a_2 \mathbf{u}_2$ is minimized.

Since the vectors $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$ and $\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$ are two solutions to the equation $zx \equiv y \pmod{2^p}$, all linear combinations of these two vectors are also solutions to this equation, in particular, $a_1 \mathbf{u}_1 + a_2 \mathbf{u}_2$ is a solution; on the other hand $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$ satisfies the equation $xz \equiv 2^q u + y \pmod{2^p}$. Hence the vector $\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} - a_1 \mathbf{u}_1 - a_2 \mathbf{u}_2$ is a solution to the equation $xz \equiv 2^q u + y \pmod{2^p}$, and it is minimal by previous result.

So, by all of the above arguments we can see that the algorithm can be used to find the minimal solution of equation (1). \square

But only find the minimal solution of equation (1) is still not enough to break the cryptosystem, because according to the system the solution should also satisfy

$$\begin{cases} 0 \leq x < 2^m \\ 0 \leq 2^q u + y < 2^m \\ 0 \leq y < 2^q, \end{cases}$$

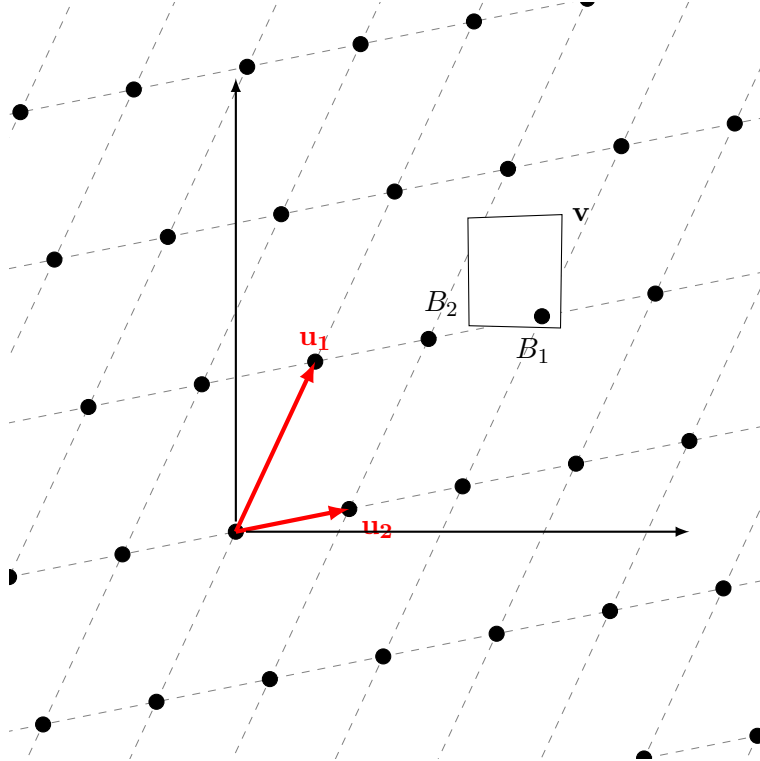
or, equivalently

$$\begin{cases} 0 \leq x < 2^m =: B_1 \\ 0 \leq y < \min\{2^q, 2^m - 2^q u\} =: B_2. \end{cases}$$

To fix our algorithm, we define an inner product on \mathbb{R}^2 by

$$\left\langle \begin{bmatrix} a_1 \\ b_1 \end{bmatrix}, \begin{bmatrix} a_2 \\ b_2 \end{bmatrix} \right\rangle = a_1 a_2 + \left(\frac{B_1}{B_2} \right)^2 b_1 b_2.$$

It is easy to see that even with this new inner product, the proof in theorem 2 is still true. Let $\{\mathbf{u}_1, \mathbf{u}_2\}$ be the minimal basis that we have found in the algorithm, and also let $\mathbf{v} = (x_0, y_0)^T$. Then we can write all four corners of the rectangle which is bounded by \mathbf{v} , $\mathbf{v} - (B_1, 0)^T$, $\mathbf{v} - (0, B_2)^T$ and $\mathbf{v} - (B_1, B_2)^T$ as real linear combinations of \mathbf{u}_1 , \mathbf{u}_2 , and then use this to find a lattice point within the bounding region (See the figure). By assumption, we know that there is at least one such lattice point; there could be more than one, but any one will solve the problem at hand.



An example based on our algorithm is as follows,

Example 1. Alice chooses her secret key $X = 12345$, which is 14-bit long, Bob and she agree on some common integers $Z = 6173$, $q = 5$ and $p = 22$, then by Bouftass's protocol, Alice needs to send

the number

$$U = \left\lfloor \frac{(XZ) \bmod(2^p)}{2^q} \right\rfloor = 708192$$

to Bob.

To recover Alice's secret key X , Eve can use the above algorithm to get:

$$x_0 = 115, \quad y_0 = 1703, \quad \mathbf{u}_1 = \begin{bmatrix} -25140 \\ 28 \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \begin{bmatrix} -33973 \\ -129 \end{bmatrix}.$$

By computing the four corners, $\mathbf{v} = 13.790\mathbf{u}_1 - 10.208\mathbf{u}_2$, $\mathbf{v} - (B_1, 0)^T = 14.252\mathbf{u}_1 - 10.108\mathbf{u}_2$, $\mathbf{v} - (0, B_2)^T = 13.531\mathbf{u}_1 - 10.016\mathbf{u}_2$ and $\mathbf{v} - (B_1, B_2)^T = 13.992\mathbf{u}_1 - 9.916\mathbf{u}_2$, Eve will find that

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 115 \\ 1703 \end{bmatrix} - 14 \begin{bmatrix} -25140 \\ 28 \end{bmatrix} + 10 \begin{bmatrix} -33973 \\ -129 \end{bmatrix} = \begin{bmatrix} 12345 \\ 21 \end{bmatrix},$$

i.e. $x = 12345$ and $y = 21$.

4 Conclusion

In this paper, we provide a practical attack to Bouftass's cryptosystem based on Gaussian lattice reduction. Our attack is simple and fast, it works when the conditions $l + m = p + q$ and $p > m + q$ are satisfied. We proved that our algorithm can definitely find a solution to the equation $u = \left\lfloor \frac{(xz) \bmod(2^p)}{2^q} \right\rfloor$, but the solution is not necessarily unique.

We also remark that a similar approach using LLL algorithm seems to work in practice, but the method presented here admitted an easier proof.

References

- [1] W. Diffie, M. Hellman, New directions in cryptography, IEEE Transactions on Information Theory 22 (6): 644–654, 1976.
- [2] Randall K. Nichols, ICSA guide to Cryptography, McGraw-Hill, New York, 1999.
- [3] S.Bouftass, On a new fast public key cryptosystem, available at arxiv.org/pdf/1508.07756.
- [4] H.Cohen, A course in computational algebraic number theory[M], Springer-Verlag: 23–24 and 83–105, 1993.
- [5] C.S.Jutla, On finding small solutions of modular multivariate polynomial equations, Advances in Cryptology-EUROCRYPT'98, Springer Berlin Heidelberg: 158-170, 1998.